A Novel Automated Biodiversity Monitoring and Conservation Information System with Google Earth Engine

Nicholas Musau Kioko, Christopher Nyaga Wanjohi, Wekesa Michael Wafula

Jomo Kenyatta University of Agriculture and Technology

Keywords: Google Earth Engine, Biodiversity, Land-Cover-Land-Use Change Detection, Earth Observation.

Abstract

Biodiversity entails a vast set of earth's life, including animals, plants, microorganisms, and organisms underwater. It contains the ecosystem variation, genetic variation, and species in the area, planet, or biome. Lately, both human and natural activities have contributed to biodiversity degradation at an alarming rate. Natural disasters like earthquakes, drought, hurricanes, and floods have significantly contributed to biodiversity disturbance. Biodiversity degradation is attracting the attention of scientists and decision-makers locally and at a global scale because of its importance in the natural reservoir and outstanding economic potential.

Locally, in Kenya, human interference in the ecosystem is greatly felt and its' effects negatively affect human life, livestock, wildlife, forests, and life underwater. The increased biodiversity degradation is associated with exponential human population growth, declining rainfall, and striking temperature rise. The economic cost of biodiversity loss in Kenya amounts to 1.3 billion dollars annually. The country's GDP is affected to the extent of 1.3 billion dollars resulting in 4.9%. More than 12 million Kenyans reside in degraded lands. Additionally, Wildlife declined by 68% threatening species extinction and population viability. Livestock numbers declined by 25.2% because of a lack of feed and sufficient water. Approximately 900000, children suffer malnutrition while adults suffer hunger stress resulting from biodiversity destruction.

Complex earth observation systems have been developed to support data and spatially extensive biodiversity research as compared to traditional approaches. Researchers intend to implement a new monitoring approach in Nakuru, Narok, and Baringo counties in Kenya considering their gradual change in biodiversity richness. The selected counties suffer increased soil erosion, agro-biodiversity loss, low productivity, forest loss, water basin reduction, and soil nutrient depletion.

This paper presents a novel approach to biodiversity conservation using remote sensing, and readily available data in Google Earth Engine (GEE) data catalog to develop a real-time monitoring biodiversity tool to classify key aspects, assess the change, and identify major contributing disturbances. Expected results are a real-time online Google Earth Engine application used for biodiversity monitoring for decision support. The researchers present a real-time biodiversity monitoring and restoration intervention tool for conservation, management, and assessment in the selected counties.

Introduction

The current state of earth surface and environmental changes warrant urgent actions from the decision makers and the civil society. Many earth observation techniques have been utilized to monitor land cover and land change detection over the years. Biodiversity entails a vast set of earth's life, including animals, plants, microorganisms, and organisms underwater. It contains the ecosystem variation, genetic variation, and species in the area, planet, or biome. Lately, both human and natural activities have contributed to biodiversity degradation at an alarming rate. Land use activities have contributed to biodiversity interference. Natural disasters like earthquakes, drought, hurricanes, and floods have significantly contributed to biodiversity disturbance. Biodiversity degradation is attracting the attention of scientists and decision-makers locally and at a global scale because of its importance in the natural reservoir and outstanding economic potential.

Objectives

The researchers propose a contemporary approach to biodiversity monitoring and conservation that focus on spatial scale. We advocate for an automated real-time monitoring approach for biodiversity conservation decision-making. The proposed and developed approach focuses on addressing the conservation gap which has existed over the years. We believe the tool will lead to increased urgent actions to help recover to functionally relevant densities. For the proposed system, we intend to develop a classification algorithm for five classes; water bodies, vegetation, grassland, and settlements. We will calculate the areas, and assess the risk extent.

Specific objectives

1. To investigate land use land cover change analysis over time for the three counties using machine learning.
2. To automate land use land cover change detection(LULCC) in google earth engine.
3. To assess land use land cover disturbances.
4. To predict the future LULC change.

Methodology

LULCC is a continuous monitoring which can detect daily changes as the Landsat get daily update with new observations. The results will place the in near real-time information with insight in the hands of decision makers. Considering the long term effect, continuous monitoring and assessment of the high frequency changes on land ought to present better understanding of the link between land use, land cover and climate variability.

Data Description

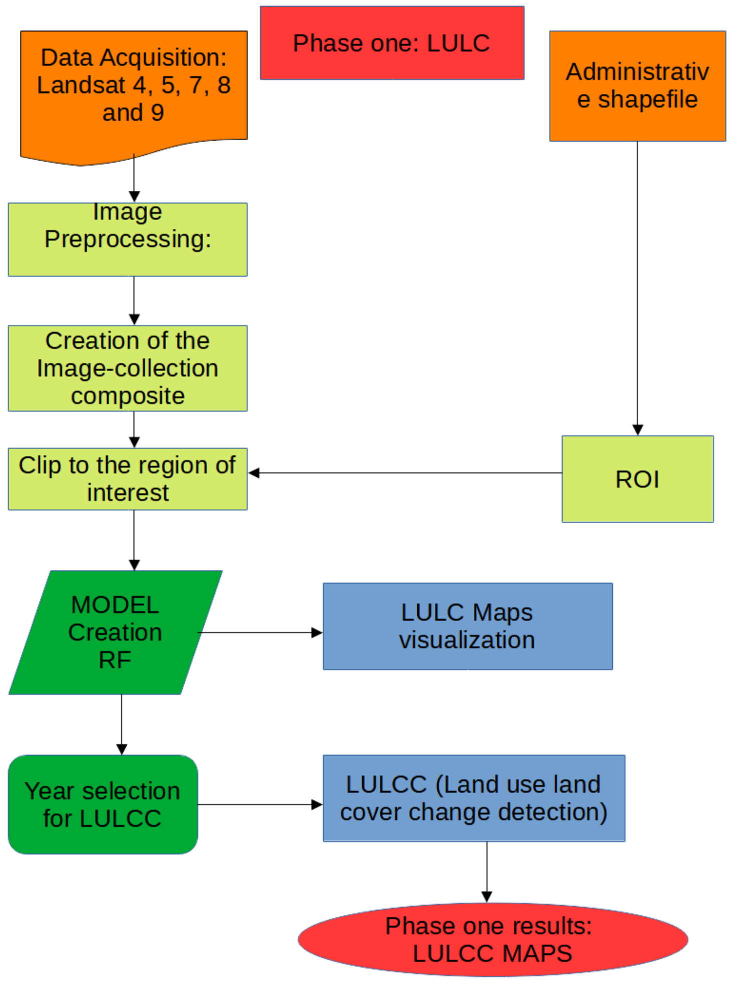
According to Sidhu et al. (2018) satellite imagery is a critical source of information in the biodiversity monitoring projects. Various sensors provide different spectral resolution aimed at detecting explicit land types. Besides the availability of varying image collections, they also vary in their spatial and temporal characteristics. Such features provide a need for clear understanding of the right sensor and understanding of the large volume of imagery dataset for better results.

Data Processing in Google Earth Engine

The research uses the free available sensor in GEE data catalog, Landsat 4, 5, 7, 8, and 9 to assess the Land Use Land Cover classification and prediction. Landsat is accessed from the United States Geological Survey. Landsat program began around 1970's which has improved over the years. According to Pimentel et al. (2014) there exist challenges when working with Landsat over the Sentinel as it had a low resolution. To counter the challenge, the sensors were fused to come up with a single highly spatially resolution imagery. The approach is commonly used as it is compatible with unsupervised algorithms which require minimal ground trothing (Mutanga & Kumar, 2019). The huge datasets were processed using the powerful GEE cloud platform which could handle large datasets at scale to assess the impact of land loss aggravators. The essence of preprocessing processes was to enhance images spectral characteristics. Resampling the image with the nearest neighbor method and pansharping panchromatic band of Landsat 7 with other bands. The next step was to reproject the subsetalgiers with a low resolution image pixel.

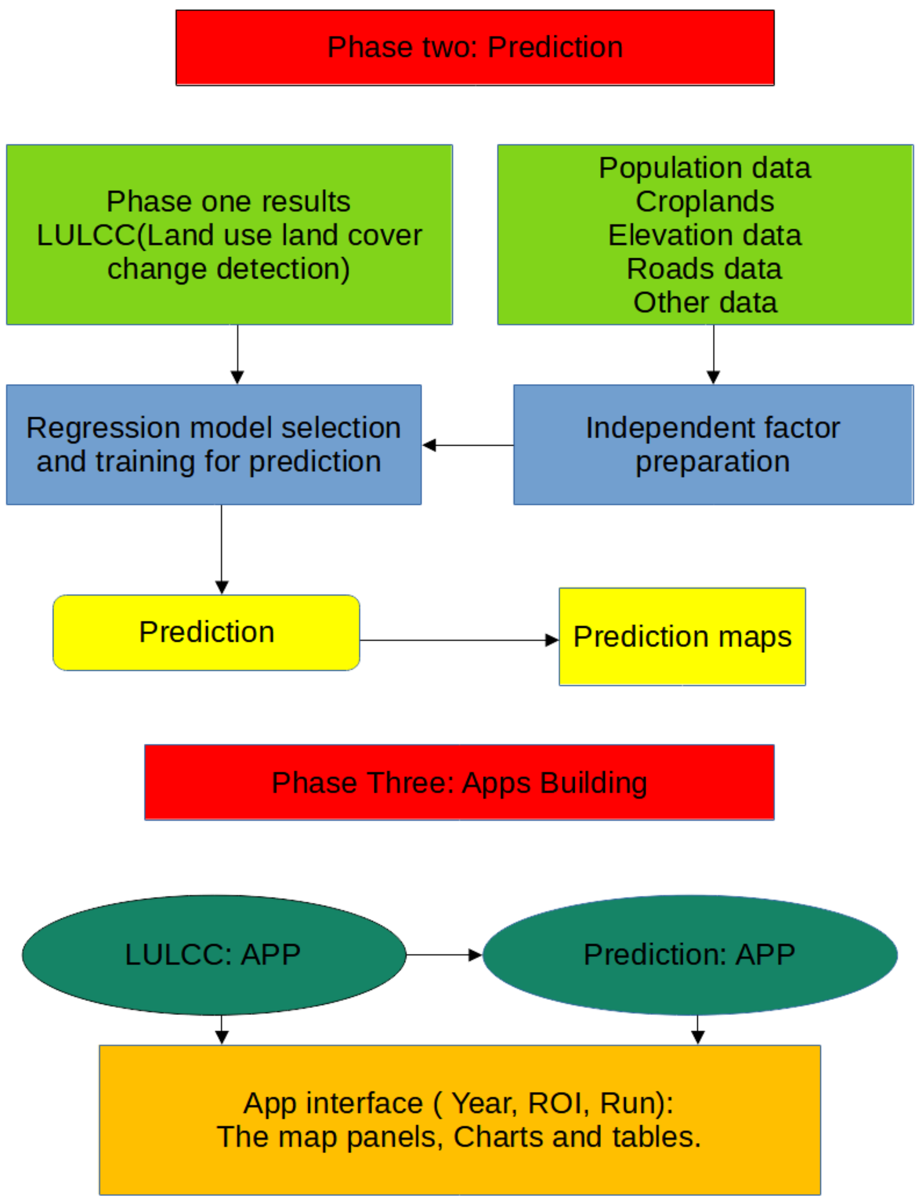
The high resolution Landsat was then moved to preprocessing stage where cloud cover was removed. The processing here led to creation of a high resolution composite. Upon acquisition of the composite, researchers clipped the image with relation to the region of interest. Clipping ensured that the classification focused on classifying LUCC within the desired study area. The next step was building the Random Forest classification model. Results of the modelling were visualized on a map. Users could select the year then on clicking an action button the visualization appear on the user interface. The preprocessing steps listed here in helped us to generate a the globally consistent analyzed data ready for classification. Classification of urban land, wet-lands, grasslands rangelands, build-up areas, non-vegetated areas mapping and agricultural lands.

Classification and Regression Tree which is a binary decision classification tree was developed for simple decision making. CART splits nodes till it reaches the terminal nodes based on the predefined thresholds. While developing the CART classifier, input data is split into sets. The trees are constructed using all the groups except one. Classification and Regression Tree's effectiveness is hampered by high dimensionally data which results in complex tree architectures. Researchers used "classifier.smileCart" technique included in the GEE JavaScript library.



During image classification, the study used six spectral bands from Landsat-8 were used. Land cover classification was divided into forest, vegetation, agricultural lands, grass ranges and buildup areas. After importing researchers removed the cloud cover and the cloud shadow. The process targeted to remove contaminated pixels from the Landsat images used in the project. The annual mean of normalized difference vegetation (NDVI) and normalized difference water indices (NDWI) were calculate within the second stage. To actualize LULC classification model, high-resolution Landsat image was used to generate 600 training polygons used for five land classes. The training was then loaded to GEE as a feature collection then NDVI and NDWI were used in the model creation.

Researchers used Random Forest classifier in the available in GEE to train Landsat composite to build ensemble classifier. Ensemble modelling entailed a combination of many classifiers in the CART trees. A range of decision trees were generated by Random Forest utilizing a random selection of training variables and dataset. Non-training sets were used to evaluate classifier's performance to provide unbiased assessment of generalization error. For establishment of an appropriate split to build a tree, Random Forest model selects variables randomly from a training sample from each node. During the model training, the most important aspect to consider are the number of trees and parameters.



Literature Review

Introduction

The land cover has over the years degraded in various parts of the world causing an alarm on the environmental status. Many contributing factors cause a dramatic change and sense of concern to the authorities, decision-makers, and communities. The unbalanced ecosystem presents threatening effects on the biomes and plants. As a result of degraded forest cover, wild animals, birds, and livestock die and lack sleeping places. According to Norris et al. (2020), the global biodiversity loss of large vertebrates from aquatic and terrestrial ecosystems has potentially a huge impact on ecosystem structure. It influences negatively, the ecological functioning, services, and biogeochemical cycles. Regulations towards land cover degradation protections have been set to safeguard and restore the lost land cover.

Land use land cover detection is a perfect approach to determine the trend of climate and human activities on the planet. Presently, the modification of the earth's terrestrial surface is a major anthropogenic factor contributing to ecosystem change. As Sidhu et al. (2018) note, the availability of technological solutions provides us with an array of options for individual targeted research. Nevertheless, the provided solutions differ based on the technological capabilities, and methodology used to develop them. Additionally, the systems vary in the classification mechanics used to generate land use classes and their distributions. Land cover detection is particularly essential in the conservation and monitoring processes. Governing authorities require improved tools for better monitoring and appropriate enforcement strategies.

Biodiversity conservation experience limitations of success because of the inability to have efficient and real-time monitoring tools for success. Additionally, a huge challenge arises in the enforcement systems because the governing bodies lack accurate argument points and risk assessment measures for proper guidelines in the governing laws. While remote sensing has for a long time shown significant importance in environmental monitoring and transformation, researchers lack the high computing power to process high volumes of satellite imagery data at scale. Additionally, the ad-hoc tools for biodiversity monitoring cannot present real-time results thus being static and with limited scalability. Publicly accessible tools like Google Earth Engine offer a considerable solution for the current and future biodiversity monitoring globally and even on a small scale Rahmi et al. (2022).

Our improved GEE tool uses publicly GEE-accessible data in the data processing stage. Further, we overcame the computational cost challenge by leveraging the GEE freely accessible computers through the GEE interface. Our solution offers actionable insights derived from daily updated satellite imagery data. Besides the existing land conservation and monitoring approaches, our BioCon App can provide insights into several environmental aspects. It presents an opportunity to utilize satellite imagery for a wide scale of applications. We could use the tool for build-up area management, water-body monitoring, agricultural land degradation, and grassland area change. More applications of the system can be integrated since it is built with scalability.

Land Cover Degradation Destructors

The issue of land cover loss is contributed by a multi-thread of factors. It adversely influences people's livelihood with high confidence. According to Gisladottir & Stocking (2005). land degradation adversely affects 25 percent of the global land area. The effects span across communities making the impact on poor people and developing nations highly hit by land degradation activities. Land use and land cover variation are direct human and climate causes. In the case of human contribution, agricultural activities dominate the land degradation aspect. Extensive soil loss resulting from the continuous land tillage exceeds the soil formation rates. Increased physical development over time which is leading to land loss affects human life greatly. It results in interference with social, political, economic, and cultural aspects. Human interference leading to land cover delineation influences the extent beyond the land surface to affect marine and freshwater systems.

Another contributing factor is climate change which exacerbates the speed and magnitude of many degradation processes. Excessive rains cause floods which lead to soil erosion. Conversely, increased temperatures lead to extensive drought which results in vegetation and grassland degradation. This leads to a lack of sufficient animal feed and loss of wild animal habitat. Degraded wetlands resulting from increased sea temperatures cause adverse effects on life underwater. Climate change cause an increased geographical area loss useful for biodiversity Norris et al. (2020).

1. **State of Earth Observation Techniques**

Biodiversity is presently a major concern in most of the political and authoritative discussion agenda. As the concept is diverse, innovative, yet scientific approaches area required to develop sustainable solutions for improved biodiversity lifestyle. Adequate technology is required for proper earth observation. Currently, satellite earth observation is a common approach to give the status of the earth. As Norris et al. (2020) notes, a number of studies succeeded in the mapping of distribution of various detailed habitat types with vegetation communities. Mapping of extensive habitat types with remote sensing techniques is a common idea reported in various land monitoring forums. Integration of Landsat datasets with field data allow the validation of the results and ascertaining of the project for implementation and decision support.Enhanced earth observation monitoring and modelling approaches help to better understand effects and potential dynamics land loss.

1. **GEE over the Existing Observation Techniques**

The Google Earth Engine cloud platform is A publicly accessible tool used to offer a considerable solution for the current and future biodiversity monitoring globally and even at small scale. While the traditional land monitoring tools are costly, Google released GEE cloud platform freely for use to researchers, developers and recently for commercial use. The existing land monitoring soft-wares are costly and challenging to learn for better use. While there are free versions, they lack the desired capabilities to produce the desired results with accuracy. GEE is easily to learn with a reliable support of experts. It is widely adopted for biodiversity conservation at a wider scale compared to other existing tools. As a result, many governing bodies are rapidly encouraging use of Google Earth Engine for real-time, monitoring of land cover.

1. **Land Change Detection Algorithms**

The problem of biodiversity degradation has faced nations for a long time. As a result, many researchers and scientist have researched and developed algorithms to help the land conservation. Good algorithms to detect land changes over time have been developed classify various aspects and the extent of delineation. Among the existing algorithms include;

Land Cover Change Algorithm

Evans et al. (2021) developed the fundamentally two algorithms to detect changes between satellite images. One of the algorithm builds on the U.S. Geological Survey system to produce the National Land Cover Dataset. The algorithm utilizes real-life looking phenomena. It calculates six spectral changes metrics for a period before and after images per pixel. It normalizes the metrics given the spread of values within the image collection. The algorithm approximates probability of change occurrence per pixel. The transformational centers and variation of pixel scales are relative to baseline changes in brightness, reflectance, and phenology between before and after imagery. The change algorithm detection for land use land cover results to a twelve bank image with normalized value and a change probability in one band.

Multivariate Alteration Detection Algorithm

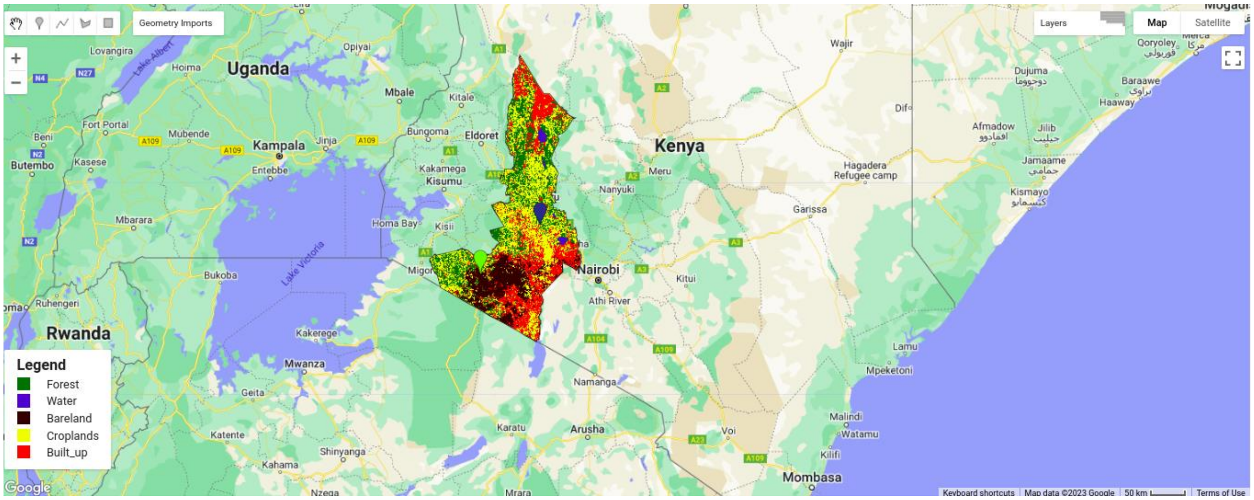
The multivariate Alteration Detection Algorithm developed by Evans et al. (2021) uses canonical correspondence analysis to identify linear transformation that use correlation between two variables of the image collection. An extreme deviation is identified per pixel through calculation of sum of the squared deviations from the mean of individual canonical variate which are relative to the variance across the image. The extreme variations represent changes in the pixel values relative to background shifts between the images. The resulting output is an image collection with several bands equivalent to the minimum number of the bands processed before and after. The bands contain values representing the canonical variates. Google Earth Engine was used in the calculation of the values and image transformation.

While there exist many land cover land use change detection, there is a gap in the automation and prediction aspect. Most of the algorithms cannot predict the expected extent of change in the future days.

Results and Discussion

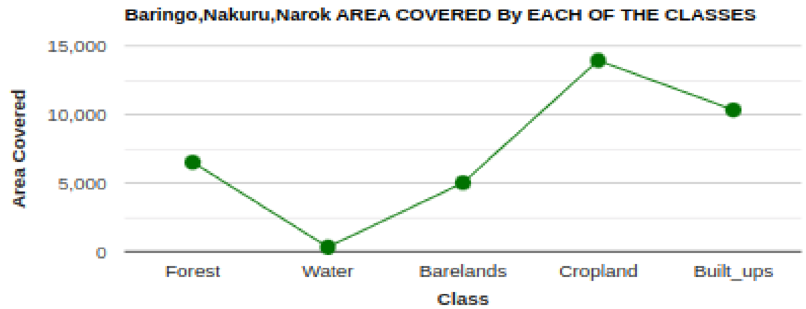
To validate our algorithm, we employed the improved change detection algorithm in three counties. We developed a real-life BioCon App that could effectively detect habitat loss, weight the loss, then predict the risk. Our approach yielded faster and efficient compared to human and existing techniques. BioCon App is online for testing and improvements are continually happening to increase its' efficiency and extent of operation. Our Google Earth Engine solution processing time varies depending on the number of years, and the internet speed. It was eventually possible to detect and measure habitat delineation in different regions resulting from both natural and anthropogenic effects. A major aspect detected is the desert loss, forest degradation from human activities, grassland coverage, and wetland reduction. Our algorithm presented a relatively high commission rates in change detecting unnoticeable changes through human inspection.

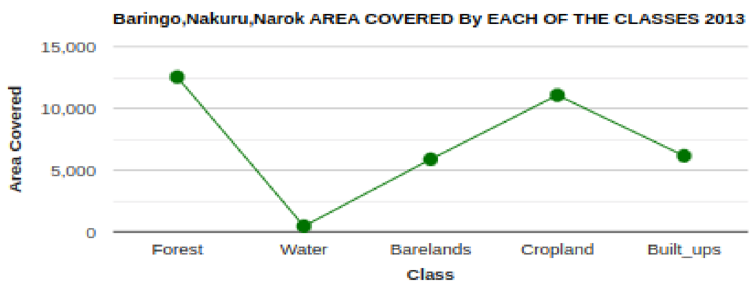
Land Cover Classification Map

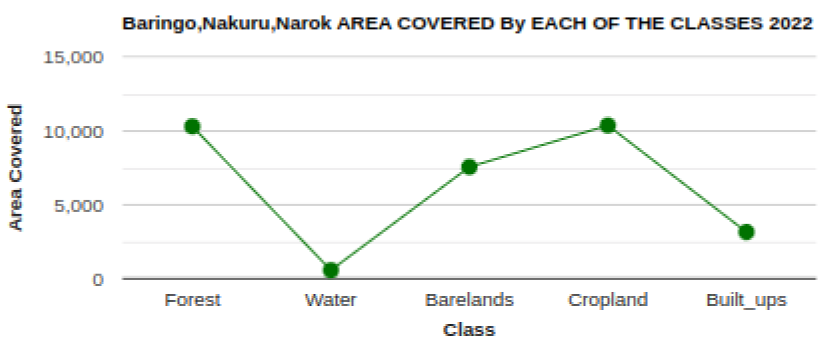


Upon carrying out spatial analysis to describe the land use and land cover change patterns the researchers arrived at the above map classification on the Google Earth Engine App. The researchers collected training data from the three counties of interest that is Baringo, Nakuru and Narok county. The map above show the classifications gained upon use of Random Forest classification in the machine model classification. We used classification data from Google Earth Engine satellite imagery then trained using 600 points per class to gain the optimum accuracy. Training samples were picked randomly and evenly in all the counties to reduce bias and enhance the model performance. The map above shows the selection of the three counties classified with the training set.

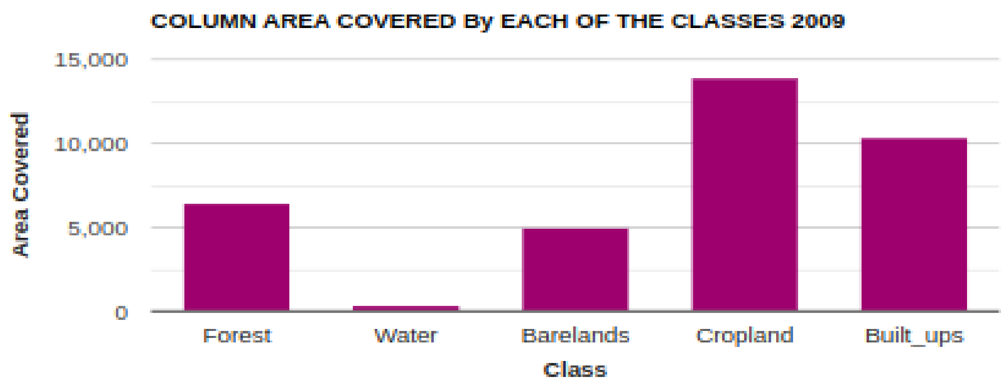
Area Coverage by Class across years

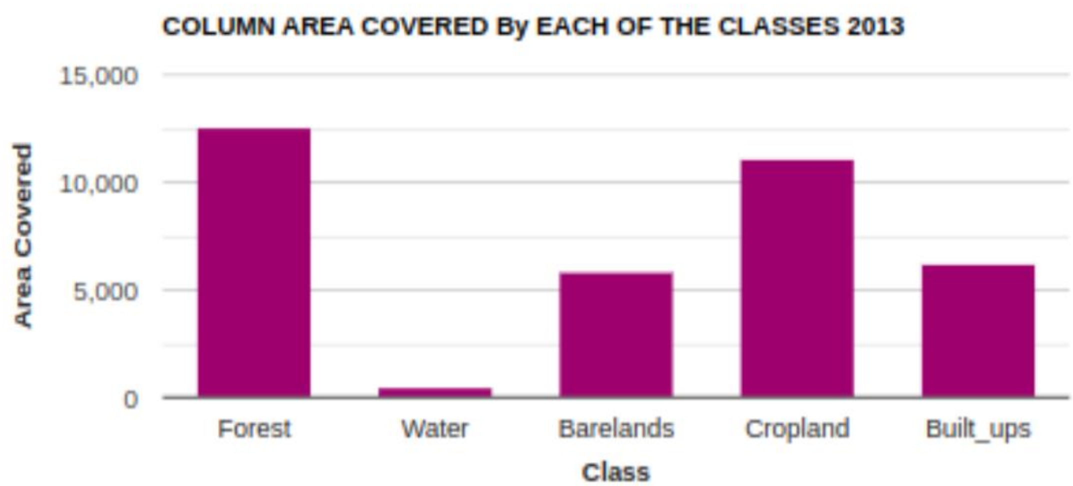


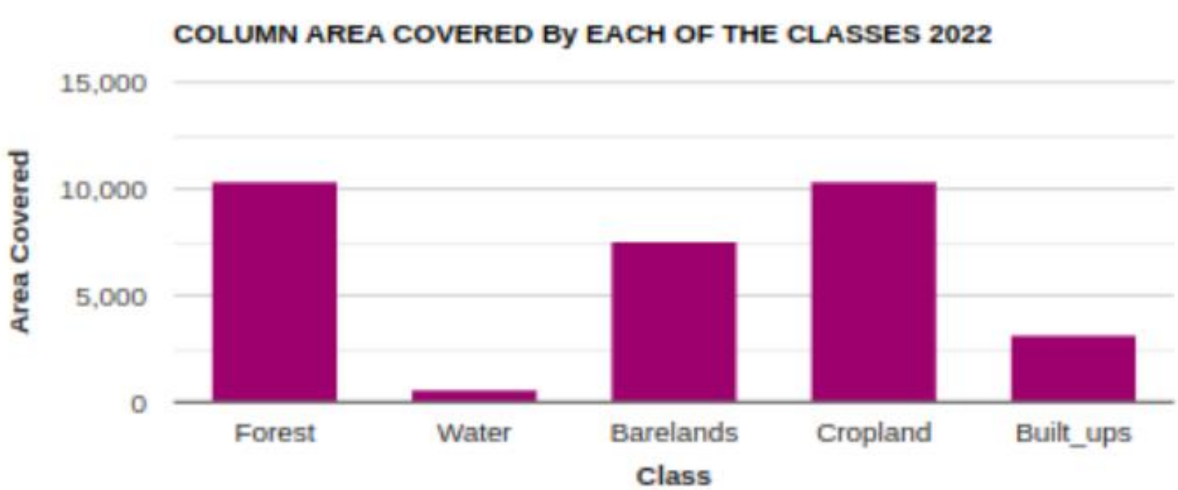




From the line charts it is evident that land use land cover change detection in the three counties presented relatively related results in terms of classification coverage. Evident results showed low water catchment regions across the counties from 2009 to 2022. The area covered by water represented approximately 1000 square kilometers. The implication of this was signaling food insecurity, livestock thirsty, degraded biodiversity. However, upon classification the researchers identified a greater land area coverage covered by crops. It is essential to note that the areas covered in the chart represented the cumulative area calculations of the three counties. Bare lands slowly increased from 2009 to 2013 and then in 2022 the area covered by bare lands increased significantly. The represented shows that drought or human activities have greatly influenced the increase of bare lands across the counties. Population growth could have also contributed to increased bare lands. People clear vegetation when constructing homesteads and preparing crop lands.







From the bar charts above shows some similarities with the line chart with regard to the land coverage across the years. Forest coverage in 2009 was about 5200 square kilometers. However, the area coverage increase gradually in 2013 to 2022 which recorded 12000 square kilometers. Forest showed the most extensive land cover classification across the years. Forested land was followed by croplands then buildup areas. While the bare land consumes a significant section of the land, agricultural activities seems to have a greater extension of the land. The variation between the two years is probably the rainfall seasons. Built up areas represented a good extension of the land coverage. It covered an estimated area of 11000 square kilometers in 2009, then reduced to 13000 in the year 2013. The buildup regions in the year 2022 reduced drastically to around 7000 square kilometers area coverage.

Recommendations

With the improved approach, land regulators and managers with or with limited knowledge in ecology can utilize our system for better and informed decisions. Policy makers with remote sensing knowledge can make wide and even inter-related solutions to serve a greater community need. Our advanced tool demonstrates efficacy, efficiency and flexibility in different counties of interest. Built on publicly and daily updated datasets and technology, the tool can be used by anyone registered in the systems and with authorization. Registered land property managers can detect urban development plus the development in the community. Local government authorities, can understand areas with urgent need of action in their strategic development and restoration plan. We perceive that the adoption and utilization of our tool can lead to improved county monitoring program. The tool automatically detects habitat degradation in real-time then assess the risk extent Meena, (2022).

The re-establishment of biome and enhanced earth observation techniques ought to become not only a global urgent priority but even locally be considered for conservation improvements and policy decisions. Researchers, scientists, developers, and policymakers ought to work collaboratively to ensure faster development of sustainable solutions for the biodiversity restoration process. Improved water and soil conservation measures ought to be established to address the low agricultural productivity. Innovative programs need to be implemented in the watershed areas.

Conclusion

Land cover land change detection if a very critical analysis in the biodiversity monitoring. It is therefore essential to carry out the classification regularly to ensure frequent results are gathered and appropriate measures taken within the shortest time possible. When the classification areas done in an automated Google Earth Engine, the local government authorities, can understand areas with urgent need of action in their strategic development and restoration plan. Researchers encourage the adoption and utilization of the Land cover land change detection tool to influence and improve county monitoring program.

References

Sidhu, N., Pebesma, E., & Câmara, G. (2018). Using Google Earth Engine to detect land cover change: Singapore as a use case. *European Journal of Remote Sensing*, *51*(1), 486-500.

Pimentel, R., Herrero, J., & Polo, M. J. (2014). Graphic user interface to preprocess Landsat TM, ETM+ and OLI images for hydrological applications.

Mutanga, O., & Kumar, L. (2019). Google earth engine applications. *Remote Sensing*, *11*(5), 591.

Norris, K., Terry, A., Hansford, J. P., & Turvey, S. T. (2020). Biodiversity conservation and the earth system: mind the gap. *Trends in Ecology & Evolution*, *35*(10), 919-926.

Rahmi, K. I. N., Ali, A., Maghribi, A. A., Aldiansyah, S., & Atiqi, R. (2022). Monitoring of land use land cover change using google earth engine in urban area: Kendari city 2000-2021. In *IOP Conference Series: Earth and Environmental Science* (Vol. 950, No. 1, p. 012081). IOP Publishing.

Evans, M. J., & Malcom, J. W. (2021). Supporting habitat conservation with automated change detection in Google Earth Engine. *Conservation Biology*, *35*(4), 1151-1161.

Meena, D. C. (2022). Land Degradation Causes, Consequences and Potential Solutions.

Gisladottir, G., & Stocking, M. (2005). Land degradation control and its global environmental benefits. *Land degradation & development*, *16*(2), 99-112.

Project Link

https://nyagachris411.users.earthengine.app/view/lulcckabarak